Rough set theory for topological spaces

E.F. Lashin [a], A.M. Kozae [b,*], A.A. Abo Khadra [a], T. Medhat [a]

[a] *Department of Physics and Engineering Mathematics, Faculty of Engineering, Tanta University, Egypt*
[b] *Department of Mathematics, Faculty of Science, Tanta University, Egypt*

Abstract

The topology induced by binary relations is used to generalize the basic rough set concepts. The suggested topological structure opens up the way for applying rich amount of topological facts and methods in the process of granular computing, in particular, the notion of topological membership functions is introduced that integrates the concept of rough and fuzzy sets. © 2005 Elsevier Inc. All rights reserved.

Keywords: Topological space; Rough set; Fuzzy set; Membership function; Knowledge representation and processing

1. Introduction

The concept of topological structures [3] and their generalizations are one of the most powerful notions in system analysis. Many works have appeared recently for example in structural analysis [4], in chemistry [17], and physics [1]. The purpose of the present work is to put a starting point for the applications of abstract topological theory into fuzzy set theory, granular computing and rough set analysis. Fuzzy set theory appeared for the first time in 1965, in famous paper by Zadeh [19]. Since

* Corresponding author.
  E-mail addresses: eflashin@yahoo.com (E.F. Lashin), akoza55@yahoo.com (A.M. Kozae), aaabokhadra@yahoo.com (A.A. Abo Khadra), tmedhat74@hotmail.com (T. Medhat).

then a lot of fuzzy mathematics have been developed and applied to uncertainty reasoning. In this theory, concepts like fuzzy set, fuzzy subset, and fuzzy equality (between two fuzzy sets) are usually depend on the concept of numerical grades of membership. On the other hand, rough set theory, introduced by Pawlak in 1982 [13], is a mathematical tool that supports also the uncertainty reasoning but qualitatively. Their relationships have been studied in [11,12,14,18]. In this paper, we will integrate these ideas in terms of concepts in topology. Topology is a branch of mathematics, whose concepts exist not only in almost all branches of mathematics, but also in many real life applications. We believe topological structure will be an important base for knowledge extraction and processing.

2. Basic concepts

Motivation for rough set theory has come from the need to represent subsets of a universe in terms of equivalence classes of a partition of that universe. The partition characterizes a topological space, called approximation space $K = (U, R)$, where U is a set called the universe and R is an equivalence relation [7,15]. The equivalence classes of R are also known as the granules, elementary sets or blocks; we will use $R_x \subseteq U$ to denote the equivalence class containing $x \in U$. In the approximation space, we consider two operators, the upper and lower approximations of subsets: Let $X \subseteq U$.

$$\overline{R}X = \{x \in U : R_x \cap X \neq \phi\},$$

$$\underline{R}X = \{x \in U : R_x \subseteq X\}.$$

Boundary, positive and negative regions are also defined:

$$BN_R(X) = \overline{R}X - \underline{R}X,$$

$$POS_R(X) = \underline{R}X,$$

$$NEG_R(X) = U - \overline{R}X.$$

These notions can be also expressed by rough membership functions [15], namely,

$$\eta_X^R(x) = \frac{|R_x \cap X|}{|R_x|}, \quad x \in U.$$

Different values defines boundary ($0 < \eta_X^R(x) < 1$), positive ($\eta_X^R(x) = 1$) and negative ($\eta_X^R(x) = 0$) regions. The membership function is a kind of conditional probability and its value can be interpreted as a degree of certainty to which x belongs to X. A quotient set version is considered in [12,10].

Fuzzy set [19] is a way to represent populations that set theory can't describe definitely, fuzzy sets use a many (usually infinite) valued membership function, unlike classical set theory which uses a two valued membership function (i.e. an element

is either in a set or it is not). Let U denotes a (universal) set and $A \subseteq U$. Then, a membership function on U, μ_A, is a function;

$$\mu_A : \ U \to L \ \text{ for some partially ordered set } L$$

L usually is a lattice [2]. Intuitively the membership function, μ_A, gives the degree to which an element $x \in U$ is in the fuzzy set A. In the case L is the closed interval [0, 1], we call it the Standard Fuzzy Set Theory.

A topological space [3] is a pair (U, τ) consisting of a set U and family τ of subset of U satisfying the following conditions:

(T1) $\Phi \in \tau$ and $U \in \tau$.
(T2) τ is closed under arbitrary union.
(T3) τ is closed under finite intersection.

The pair (U, τ) is called a space, the elements of U are called points of the space, the subsets of U belonging to are called open set in the space, and the complement of the subsets of U belonging to τ are called closed set in the space; the family τ of open subsets of U is also called a topology for U.

It often happens that the open sets of space can be very complicated and yet they can all be described using a selection of fairly simple special ones. When this happens, the set of simple open sets is called a base or subbase (depending on how the description is to done). In addition, it is fortunate that many topological concepts can be characterized in terms of these simpler base or subbase elements. Formally, A family $\beta \subseteq \tau$ is called a base for (U, τ) iff every non_empty open subset of U can be represented as a union of subfamily of β. Clearly, a topological space can have many bases. A family $S \subseteq \tau$ is called a subbase iff the family of all finite intersections is a base for (U, τ).

$\overline{A} = \cap \{F \subseteq U : A \subseteq F \text{ and } F \text{ is closed}\}$ is called the τ-closure of a subset $A \subseteq U$.

Evidently, $\overline{A}$ is the smallest closed subset of U which contains A. Note that A is closed iff $A = \overline{A}$.

$A^\circ = \cup \{G \subseteq U : G \subseteq A \text{ and } G \text{ is open}\}$ is called the τ-interior of a subset $A \subseteq U$.

Evidently, A° is the union of all open subsets of U which containing in A. Note that A is open iff $A = A^\circ$. And

$A^b = \overline{A} - A^\circ$ is called the τ-boundary of a subset $A \subseteq U$.

3. Rough set theory in topological spaces

The reference space in rough set theory is the approximation space whose topology is generated by the equivalence classes of R. This topology belongs to a special class known by Clopen topology, in which every open set is closed. Clopen topology

is called quasi-discrete topology in digital geometry; Lin calls it Pawlak space [7]. Clopen topology is a kind of approximations that are transitive. It is too restrictive, for example, "East LA is close to LA, LA is close to West LA. However, East LA is not considered to be close to West LA" [7]; approximations are often not transitive. Lin introduced neighborhood system to handle such general situations [5,6,8]. We will use topology; in other words, the "approximation space" is a topological space.

We will express rough set properties in terms of topological concepts. Let X a subset. Let $\overline{X}, X^\circ$ and X^b be closure, interior, and boundary points respectively. X is exact if $X^b = \Phi$, otherwise X is rough. It is clear X is exact iff $\overline{X} = X^\circ$. In Pawlak space a subset $X \subseteq U$ has two possibilities rough or exact. For a general topological space, $X \subseteq U$. X has the following types of definability:

(1) X is totally definable if X is exact set "$\overline{X} = X = X^\circ$",
(2) X is internally definable if $X = X^\circ$, $X \neq \overline{X}$,
(3) X is externally definable if $X \neq X^\circ$, $X = \overline{X}$,
(4) X is undefinable if $X \neq X^\circ$, $X \neq \overline{X}$.

Proposition 1. *If A is an exact set in (U, τ) and $\tau \subset \tau'$ then A is exact with respect to τ'.*

Proof. Since $BND_\tau A \subset BND_\tau A$ and $BND_\tau\, A = \Phi$. Then $BND_{\tau'} A = \Phi$ and A is exact with respect to τ'. In other words if A is τ_exact then A is τ_clopen and consequently τ'_clopen. Hence A is τ'_exact. $\square$

It is easy to have examples for a τ'_exact set which is not τ_exact. Let us observe that $cl_{\tau'}A = cl_\tau A$ iff $int_{\tau'}A^c = int_\tau A^c$. The following proposition gives the condition for τ'_exact sets to be τ_exact sets, $\tau \subset \tau'$.

Proposition 2. *If (U, τ) is a space and $\tau \subset \tau'$ then each exact set in τ' is exact in τ' iff $cl_\tau G = cl_{\tau'}G, \quad \forall G \in \tau'$.*

Proof. If A is τ'_exact then $cl_{\tau'}A = A$ and $cl_\tau A = A$, hence $cl_\tau A = cl_{\tau'}A$. Conversely: if $cl_\tau A = cl_{\tau'}A$ and A is τ'_exact. Then A is τ_exact. $\square$

Original rough membership function is defined using equivalence classes. We will extend it to topological spaces. If τ is a topology on a finite set U, where its base is β, then the rough membership function is

$$\mu_X^\tau(x) = \frac{|\{\cap B_x\} \cap X|}{|\cap B_x|}, \quad B_x \in \beta, \ x \in U$$

where B_x is any member of β containing x. It can be shown that this number is independent of the choice of bases. Since, the intersection of all members of the topology containing x concedes with the intersection of all members of a base containing x.

Note that if the topology is the Clopen one x belongs to a unique member of the base. Moreover the above membership function give the ordinary set theory if τ is discrete topology and rough set theory if τ is Clopen(quasi discrete) topology.

The following example illustrates the above definition. Let $U = \{0,1,2,3,4,5\}$, $\beta = \{\{2\},\{3\},\{0,1,2\},\{2,3,4\}\{3,5\}\}$, $X = \{2,4,5\}$, we get:

$$\mu_X^\tau(0) = \frac{|\{0,1,2\} \cap \{2,4,5\}|}{|\{0,1,2\}|} = 1/3, \quad \mu_X^\tau(3) = 0,$$

$$\mu_X^\tau(1) = 1/3, \qquad\qquad\qquad \mu_X^\tau(4) = 2/3,$$

$$\mu_X^\tau(2) = 1, \qquad\qquad\qquad\quad \mu_X^\tau(5) = 1/2.$$

In the case of infinite universe, this membership function can be use for spaces having locally finite neighborhood systems in the sense that there are only finitely many minimal neighborhoods for each point.

Rough membership functions allow us to express fuzzy theory in topological spaces: Let $X \subseteq U$ be a subset, we define a fuzzy set by using the rough membership function of topological spaces

$$\underset{\sim}{X} = \{(x, \mu_X^\tau(x)) : \forall x \in U\}.$$

From the above example, we find that: if $x = \{2,4,5\}$.
Then $\underset{\sim}{X} = \{(0,1/3), (1,1/3), (2,1), (3,0), (4,2/3), (5,1/2)\}$.

4. Rough set theory in the topology of binary relation

As we have pointed out earlier that Lin introduced the formalism of neighborhood system to handle such general situations. We will consider the topology generated from the binary relation R. If U is a finite universe and R is a binary relation on U, then we define, right neighborhood

$$xR = \{y : xRy\}$$

We should note that xR is a right neighborhood of x, but xR is *not* necessary a right neighborhood of any element in xR. In fact, the set of all elements, each of which has xR as its right neighborhood, is called the center of xR. The collections of all centers form a partition of U; see [8] for details.

We will not consider right neighborhood system (T.Y. Lin skips the word right), we will consider the topology generated by right neighborhoods. Taking such view xR is an open set, which is a neighborhood (in the sense of topological space) of each of its points. To construct the topology, we consider the family $S = \{xR : x \in U\}$ of right neighborhood as a subbase. Let the induced topology be τ. The family S as the subbase of τ will be denoted by $S_R = \{xR : x \in U\}$, and we write $S_x = \{G \in S_R : x \in G\}$.

Since all finite intersections of members of a subbase form a base, the notion of topological rough membership functions can be expressed by subbase:

$$\mu_X^\tau(x) = \frac{|\{\cap S_x\} \cap X|}{|\cap S_x|}, \quad x \in S_x, \ S_x \in S$$

Note that this rough membership is very different from rough set theory or Lin's rough membership function of right neighborhood. In Lin's case instead of $\cap Sx$, he will use xR, which is unique; we will report the difference in future work. It may exist $y \in U$ and y belongs to more than one S_x as shown in the following example; note that one of S_x and xR is the same as sets. However, they are different xR is a right neighborhood and is unique (in the formalism of Lin's neighborhood system), while S_x is a set of open neighborhood of x in the topology τ.

Example 1. Let $U = \{0, 1, 2, 3, 4, 5\}$, $0R = 1R = \{0, 1, 2\}$, $2R = 3R = \{2, 3\}$, $4R = \{3, 4\}$, $5R = \{5\}$ Then $S := \{\{0, 1, 2\}, \{2, 3\}, \{3, 4\}, \{5\}\} \Rightarrow \beta = \{\{0, 1, 2, \}, \{2, 3\}, \{3, 4\}, \{5\}, \{2\}, \{3\}\} \Rightarrow \tau = \{U, \Phi, \{0, 1, 2\}, \{2, 3\}, \{3, 4\}, \{5\}, \{2\}, \{3\}, \{0, 1, 2, 3\}, \{0, 1, 2, 3, 4\}, \{0, 1, 2, 5\}, \{2, 3, 4\}, \{2, 3, 5\}, \{3, 4, 5\}, \{2, 5\}, \{3, 5\}, \{2, 3, 4, 5\}\}$.

Let $X = \{0, 1, 2, 3\}$

$$\therefore \mu_X^\tau(0) = \frac{|\{0, 1, 2\} \cap \{0, 1, 2, 3\}\|}{|\{0, 1, 2\}|} = 1, \quad \mu_X^\tau(3) = 1,$$

$$\mu_X^\tau(1) = 1, \qquad\qquad\qquad\qquad \mu_X^\tau(4) = 1/2,$$

$$\mu_X^\tau(2) = 1, \qquad\qquad\qquad\qquad \mu_X^\tau(5) = 0.$$

Then $\underset{\sim}{X} = \{(0, 1), (1, 1), (2, 1), (3, 1), (4, 1/2), (5, 0)\}$

From rough membership function, we get:

$$\underline{R}X = X^\circ = \{0, 1, 2, 3\}, \quad \overline{R}X = \overline{X} = \{0, 1, 2, 3, 4\},$$

$$NEG_R(X) = \{5\}, \quad BN_R(X) = \{4\}$$

We can get the interior and closure of X by using the definitions of τ-closure and τ-interior without using the membership function as follows: Here are the family F of all τ-closed sets:

$$F = \{\Phi, U, \{3, 4, 5\}, \{0, 1, 4, 5\}, \{0, 1, 2, 5\}, \{0, 1, 2, 3, 4\}, \{0, 1, 3, 4, 5\},$$
$$\{0, 1, 2, 4, 5\}, \{4, 5\}, \{5\}, \{3, 4\}, \{0, 1, 5\}, \{0, 1, 4\}, \{0, 1, 2\}, \{0, 1, 3, 4\},$$
$$\{0, 1, 2, 4\}, \{0, 1\}\}.$$

So,

$$X^\circ = \{0, 1, 2\} \cup \{2, 3\} \cup \{2\} \cup \{3\} = \{0, 1, 2, 3\},$$

$$\overline{X} = U \cap \{0, 1, 2, 3, 4\} = \{0, 1, 2, 3, 4\}.$$

5. Granular structure in the topology of binary relations

The purpose of this section is to investigate the knowledge representations and processing of binary relations in the style of rough set theory. Let us consider the

pair, (U, B), where $B = \{R1, R2, \ldots, Rn\}$ is a family of general binary relations on the universe U. When B is a family of equivalence relations, Pawlak call it knowledge base and Lin call the general case binary knowledge base in [8]. As the term "knowledge base" often means something else, Lin begin to use the generic name granular structure [8,9]. We will use knowledge structure and granular structure interchangeably.

Next, we will consider the topological space for each binary relation; we will call it the topological space of the binary relation (TSB). We denote the base β_R that is generated by the binary relation R. Note that two distinct binary relations R and R' may generate the same topology as shown in the following example: Let $U = \{0, 1, 2, 3, 4, 5\}$, R and R' are distinct binary relations, where

$$R = \{(0,0), (0,1), (0,2), (1,2), (1,3), (2,2), (2,3), (3,2), (3,3), (4,3), (4,4), (5,5)\}$$

$$R' = \{(0,0), (0,1), (0,2), (1,0), (1,1), (1,2), (2,2), (2,3), (3,3), (3,4), (4,3),$$
$$(4,4), (5,5)\}$$

Their (right) neighborhood systems are: (as subbases)

$$0R = \{0, 1, 2\}, 1R = 2R = 3R = \{2, 3\}, \quad 4R = \{3, 4\}, 5R = \{5\},$$
$$0R' = 1R' = \{0, 1, 2\}, \quad 2R' = \{2, 3\}, \quad 3R' = 4R' = \{3, 4\}, \quad 5R' = \{5\}.$$

These two subbases generated the same base $S_R = \{\{0,1,2\}, \{2,3\}, \{3,4\}, \{5\}\} = S_{R'}$, hence the same topology $\tau_R = \tau_{R'}$.

Next, we will generalize the notion of reducts to TSB, the topological space of binary relations.

Definition 1. Let $P \subseteq B$ be a subset of B, $r \in P$, where B be a class of binary relations. r is said to be superfluous binary relation in P if:

$$\beta_P = \beta_{(P-\{r\})}$$

The set M is called a minimal reduct of P iff:

(i) $\beta_M = \beta_{(P)}$.
(ii) $\beta_M \neq \beta_{(P-\{r\})}, \ \forall r \in M$.

The following example illustrates the notion given above.

Example 2. Let $U = \{1, 2, 3, 4, 5\}$, and the 3 subbases $S_r = \{\{1,2\}, \{2,3,4\}, \{4,5\}\}$, $S_p = \{\{1,2,3\}, \{3,4\}, \{5\}\}$, $S_q = \{\{1,2\}, \{3,4\}, \{4,5\}\}$. Then we have a joint subbase

$$S_B = \{\{1,2\}, \{2,3,4\}, \{4,5\}, \{1,2,3\}, \{3,4\}, \{5\}\}.$$

The base is $\beta_B = \{\{1,2\}, \{2\}, \{3\}, \{4\}, \{5\}\}$, Next consider

$$S_{(B-r)} = \{\{1,2,3\}, \{3,4\}, \{5\}, \{1,2\}, \{4,5\}\},$$
$$\beta_{(B-r)} = \{\{1,2\}, \{3\}, \{4\}, \{5\}\},$$

$$S_{(B-p)} = \{\{1,2\},\{2,3,4\},\{4,5\},\{3,4\}\},$$
$$\beta_{(B-p)} = \{\{1,2\},\{2\},\{3,4\},\{4\},\{4,5\}\},$$

$$S_{(B-q)} = \{\{1,2\},\{2,3,4\},\{4,5\},\{1,2,3\},\{3,4\},\{5\}\},$$
$$\beta_{(B-q)} = \{\{1,2\},\{2\},\{3\},\{4\},\{5\}\} = \beta_B.$$

So we find that q is only superfluous relation in B, and we have

$$\text{RED}(B) = \{r,p\}, \quad \text{CORE}(B) = \{r,p\}.$$

6. Conclusions

In this work, we generalize rough set theory in the frameworks of topological spaces. We believe such generalized rough set theory will be useful in digital topology [16] as well as biomathematics [17]. Our approach in essence is to topologize information tables (also known as information systems). Our theory connects rough sets, topological spaces, fuzzy sets, and neighborhood systems (binary relations, pretopology). This theory brings in all these techniques to information analysis and knowledge processing. We believe that topological structure is the appropriate 'umbrella'.

Acknowledgment

References

[1] A. Galton, A generalized topological view of motion in discrete space, Theor. Comput. Sci. (2002).
[2] M. Goldstern, Lattices, interpolation and set theory, In Contr. General Algebra 12 (2000) 23–36.
[3] J. Kelley, General Topology, Van Nostrand Company, 1955.
[4] C. Largeron, S. Bonnevay, A Pretopological Approach for Structural Analysis, Information Sci. 144 (2002) 185–196.
[5] T.Y. Lin, Neighborhood systems and relational database, in Proceedings of 1988 ACM Sixteen Annual Computer Science Conference, February 23–25 (1988) 725.
[6] T.Y. Lin, K.J. Huang, Q. Liu, and W. Chen, Rough Sets, Neighborhood systems and approximation, in: Proceedings of the Fifth International Symposium on Methodologies of Intelligent Systems, Selected Papers, Knoxville, Tennessee, October 25–27 (1990) 130–141.
[7] T.Y. Lin, Topological and fuzzy rough sets, in: R. Slowinski (Ed.), Decision Support by Experience—Application of the Rough Sets Theory, Kluwer Academic Publishers, 1992, pp. 287–304.
[8] T.Y. Lin, Granular computing on binary relation I: Data mining and neighborhood systems, in: Rough Sets In Knowledge Discovery, Physica-Verlag, 1998, pp. 107–121.
[9] T.Y. Lin, Granular computing on binary relations II: Rough set representations and belief functions, in: A. Skowron, L. Polkowski (Eds.), Rough Sets In Knowledge Discovery, Physica-Verlag, 1998, pp. 121–140.

[10] T.Y. Lin, Sets with partial memberships: A rough sets view of fuzzy sets, World Congress of Computational Intelligence (WCCI98), Anchorage, Alaska, May 4–9 (1998) 785–790.

[11] T.Y. Lin, Granular computing: Fuzzy logic and rough sets, in: L.A. Zadeh, J. Kacprzyk (Eds.), Computing with words in information/intelligent systems, Physica-Verlag (A Springer-Verlag Company), 1999, pp. 183–200.

[12] T.Y. Lin, Granular fuzzy sets: A view from rough set and probability theories, Int. J. Fuzzy System 3 (2) (2001) 373–381.

[13] Z. Pawlak, Rough sets, Int. J. Information Comput. Sci. 11 (5) (1982) 341–356.

[14] Z. Pawlak, Rough sets and fuzzy sets, Fuzzy Sets And Systems 17 (1985) 99–102.

[15] Z. Pawlak, Rough Sets, Theoretical Aspects of Reasoning about Data, Kluwer Academic, Boston, 1991.

[16] A. Rosenfeld, Digital topology, American Mathematical Monthly 86 (1979) 621–630.

[17] B.M.R. Stadler, P.F. Stadler, Generalized topological spaces in evolutionary theory and combinatorial chemistry, J. Chem. Inf. Comput. Sci. 42 (2002) 577–585.

[18] M. Wygralak, Rough sets and fuzzy sets some remarks on interrelations, Fuzzy Sets And Systems 29 (2) (1989) 241–243.

[19] L.A. Zadeh, Fuzzy sets, Information and Control 8 (1965) 338–352.